INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Fast Non Rigid Matching by Gradient Descent: Study and Improvements of the "Demons" Algorithm

Pascal Cachier, Xavier Pennec, Nicholas Ayache

N° 3706

June 1999

THÈME 3

ISSN 0249-6399

INRIA
SOPHIA ANTIPOLIS

Fast Non Rigid Matching by Gradient Descent: Study and Improvements of the "Demons" Algorithm

Pascal Cachier, Xavier Pennec, Nicholas Ayache

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projet Epidaure

Rapport de recherche n° 3706 — June 1999 — 25 pages

Abstract: Most iconic methods for rigid matching consist in finding and minimizing a registration criterion specifically chosen to solve a given problem. For non-rigid matching, attention has rather focussed on the type of smoothing or physical model of deformation to be used. In this report, we propose to place the non-rigid matching problem into a minimization framework. We have developped our theoretical idea in the case of the least squares criterion, corresponding to the assumption that the intensities of points do not change over time, and we have implemented a first order gradient descent which, along with a multiresolution approach, minimizes this criterion. We also prove that the "demons" algorithm, thought of until now as an *as hoc* matching technique, can be seen as an approximation of a second order gradient descent on this criterion.

Analysis of the mechanisms of this gradient descent incites us to introduce two different weightings into the filters used to smooth the solution, which we called an *a priori* weighting and an *a posteriori* weighting. These weightings, which are independent of the filter, have both proven their efficiency in their own way: the *a priori* weighting improves the solution found for the minimization problem, which is shown by comparing results in a distance-roughness space, while the *a posteriori* weighting helps tackle the appearance or disappearance of matter and occlusions, both sensitive issues for non-rigid iconic methods.

Key-words: Non-rigid matching, least squares method, gradient descent, "demons" algorithm, filter weighting, distance-roughness space.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, B.P. 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : 04 92 38 77 77 - International : +33 4 92 38 77 77 — Fax : 04 92 38 77 65 - International : +33 4 92 38 77 65

Recalage Non Rigide Rapide par Descente de gradient: Etude et Amélioration de l'Algorithme des Démons

Résumé : En recalage rigide, la majorité des méthodes iconiques consistent à trouver et minimiser un critère de recalage adapté au problème. En recalage non rigide, l'intérêt a plutôt porté sur le type de lissage effectué, voire le modèle physique de transformation utilisé. Nous proposons dans ce rapport de replacer le problème du recalage non rigide dans un formalisme de minimisation. Nous développons pour cela les calculs dans le cas du critère des moindres carrés, correspondant à l'hypothèse simple de conservation de l'intensité d'un point au cours du temps, et nous proposons une descente de gradient couplée à une approche multiéchelle pour résoudre cette minimisation. Nous montrons au passage que l'algorithme des "démons", technique de recalage jusqu'à présent *ad hoc*, peut être considéré comme une approximation d'une descente de gradient du second ordre sur ce critère.

L'analyse de la descente de gradient effectuée nous amène à introduire deux pondérations dans les filtres utilisés pour régulariser la solution, les pondérations *a priori* et *a posteriori*. Ces pondérations, indépendantes du filtre, ont chacune prouvé leur efficacité dans notre implémentation: la pondération *a priori* en améliorant l'efficacité et la vitesse de la minimisation, ce que nous montrons en étudiant les résultats de recalage dans un espace distance-rugosité, et la pondération *a posteriori* en gérant les problèmes d'apparition et de disparition de matière, problèmes particulièrement délicats pour les méthodes iconiques non rigides.

Mots-clés : Recalage non rigide, méthode des moindres carrés, descente de gradient, algorithme des démons, filtres pondérés, espace distance-rugosité

1 Introduction

Image matching consists in retrieving the position of each point of an object in a sequence of images depicting this object. This problem is widely spread in many fields of image processing, and its applications are numerous, e.g., motion tracking and video compression. Image matching is particularly useful in medical imaging, as it has many applications: construction of anatomic atlas, detection of cerebral lesion evolution, tracking surgical tools in images taken during an operation, etc. Formalizing a little bit, we could say that matching two images I and J amounts to retrieving the transformation T, defined in a plane or in space according to the dimension of images, such that $I = J \circ T$.

Depending on the space to which the transformation T belongs, we can classify matching algorithms in two classical categories. On one hand, the *rigid* or *affine* matching algorithms let T have only a few degrees of freedom. On the other hand, the *non-rigid* matching algorithms look for T in a very high dimensional space. Most of the time, the choice between either a rigid or a non-rigid matching algorithm is dictated by the medical problem considered; for example, a rigid method will be used to match T1 MR and CT images of the head of a patient, whereas brain detection of lesion evolution in a set of MR images taken weeks apart from each other require a non-rigid matching algorithm. Note, however, that most of non-rigid algorithms need initial rigid alignment.

Lots of studies have already been done on registration [MV98]. For example, some methods use external markers, such as stereotaxic frames [LJ94], which have thus far been considered as the gold standard for rigid registration, others use anatomical landmarks, such as crest lines extracted from the images[Sub98]. However, most of these techniques are used for rigid matching. Non-rigid algorithms usually work directly with the intensity of pixels, because they need denser information everywhere in the images, and because the geometric landmarks are usually no longer invariant over non rigid transformations. Such algorithms are said to be iconic.

The difference in the way rigid and non-rigid iconic methods are dealt with is very surprising. For rigid matching, researchers have produce some criteria measuring the apparent similarity of two images. These criteria have been compared in numerous papers [RMPA98, SHH95]; among them, let us mention the correlation coefficient [ORSP98], entropy of joint histogram [CVSM95], mutual information [Vio95] and correlation ratio [RMPA98]. Rigid registration consists in minimizing this criterion among the set of rigid transformations.

For non-rigid matching, however, most authors are not looking to minimize a criterion, but rather study and compare the different kinds of deformation smoothing, which has incidentally given birth to some new terminology, including "elastic matching" [Bro81, BK89] and "viscous fluid matching"[CJM97, BNG96].

In this report, we propose to put the non-rigid matching problem into a criterion minimization framework and to use a tool adapted to the minimization of a function with many parameters: the gradient descent [GMW81]. Then, we introduce two new weights in the transformation filters, the *a priori* and the *a posteriori* weights, which respectively help the gradient descent to find a better solution and tackle typical iconic method problems regarding the appearance or disappearance of matter and occlusions. Finally, to demonstrate the improvements done, we propose to compare non-rigid matching algorithms in an image distance/transformation roughness space.

2 3D Image Matching

2.1 Quick Overview of Non-Rigid Matching

Up to our knowledge, Broit [Bro81] and Bajcsy [BK89] were the first to publish on non-rigid matching. Broit used correlation

$$C(f,g) = \frac{< f,g >}{(< f,f >< g,g >)^{1/2}} \qquad \text{with } < f,g >= \int f.g$$

as a local similarity measure between two images. Bajcsy differentiated this criterion and used a fixed fraction of its gradient as external forces for a linear elastic model describing the behavior of deformations. Bajcsy was also the first to introduce the multiscale approach in her non-rigid matching algorithm.

Christensen [CJM97] proved that linear elastic models for transformations do not guarantee the preservation of the image topology for large motions. More precisely, the determinant of the Jacobian of the transformation could become negative. Thus, he proposed to use a viscous fluid model for transformations instead, arguing that this model can handle large motions without breaking the image topology. In practice, however, this model is also linearized.

Bro-Nielsen [BNG96] continued working on Christensen model, and used the fact that the partial derivative equations are linear to deduce the corresponding linear filters that advantageously replace the finite element resolution of Christensen; his algorithm thus gained several orders of magnitude in the computation speed. Furthermore, Bro-Nielsen presented his forces as being the differential of a sum of square differences (SSD) criterion; however, he decided to use a fixed fraction of this differential.

Thirion [Thi98] proposed to look at non-rigid matching as a diffusion process. From this viewpoint, the author decided to introduce into the image some entities which he called "demons", by analogy with Maxwell's demons introduced to explain Gibbs' paradox in thermodynamics. Thirion gave several methods using "demons", depending on whether images are segmented or not. For non-segmented images, his technique, explained in section 3 below, starts from the equations of optical flow, which has also been applied to non-rigid matching by several other authors [VHSL98].

Lastly, some authors have applied rigid matching criteria to non-rigid matching, using a hybrid scheme called block matching [MMV98, WKJK98]. However, since the block usually cannot have a small size for robustness reasons, the resolution of this method is *a priori* limited. Furthermore, in the case where the position of blocks are found by exhaustive translational search over a region, local rotations may hamper the algorithm and can be difficult to retrieve, and also computation times can be long for 3D images.

2.2 Formalization

In this report, an image I is defined to be a function of $[0;1]^n$ into $\mathbf{R}$, with $n = 2$ or 3. For a particular point $\mathbf{p} \in [0;1]^n$, $I(\mathbf{p})$ is the value of the intensity at $\mathbf{p}$. This number corresponds to the value of the physical property measured by the modality of I around this point, like the X-ray absorption density or the proton emission density. Furthermore, this value may be either directly measured or retrieved by tomography. For a good overview of the medical imaging physics, see [BT88].

Once we are given two images I and J, a set of 3D transformations $\mathcal{T}$, and a metric d between images, matching consists in solving the following minimization problem, find $T^* \in \mathcal{T}$ such that

$$T^* = \arg\min_{T \in \mathcal{T}} d(I, J \circ T) \tag{1}$$

It is quite important to notice that the minimization problem (1), due to its formulation and due to the methods used to solve it, is dissymmetric with respect to I and J, while the matching problem is *a priori* symmetric.

The first reason for this dissymmetry comes from the fact that T is not necessarily invertible. Smoothings try to avoid this unrealistic case.

When we suppose that T is invertible, the main reason for the dissymmetry is the metric: $d(I, J \circ T) \neq d(I \circ T^{-1}, J)$. This implies that we will have a different matching solution when we invert the role of I and J.

Even if the metric d is symmetric, another cause of dissymmetry is introduced by minimization algorithms used to solve the problem. For example, a gradient descent will not have the same direction of search if we invert I and J. Even if the global minimum is the same, the algorithm may be trapped in different local minima.

In the case of non-rigid matching, the last cause of dissymmetry comes from the various smoothings done on the transformation T, which usually do not commute with inversion, i.e. if we apply a smoothing S on T then:

$$[S(T)]^{-1} \neq S(T^{-1})$$

This is the case for the classical linear elastic and Gaussian smoothings. Transformation smoothing is an unavoidable step for non-rigid matching, in order both to regularize the problem (1), and to improve the fit of the transformation T to the real transformation that has occured and which is subject to physical constraints. Let us highlight the fact that the transformation T is defined over the points of the image I and is constructed so that its values are the positions of the corresponding points in the image J, i.e., $\mathbf{p}$ and $T(\mathbf{p})$ are supposed to be the coordinates of the same physical point respectively in I and in J. Motion has occurred for this point if $\mathbf{p} \neq T(\mathbf{p})$. Be careful, however, as the image I deformed by transformation T is $I \circ T^{-1}$, not $I \circ T$. In practice, it is more convenient to compare I and $J \circ T$ instead of $I \circ T^{-1}$ and J, for we do not need to invert the transformation.

Because I and J play different roles, we introduce some terminology to distinguish between them. The image I is called the reference image and J the study image. Because the result of our algorithm is $(I, J \circ T)$, as explained above, J is deformed while I is not and so this terminology makes sense.

The images are only made of values for a discrete set of points in space. Therefore, we have to interpolate them in order to obtain values over the entire space. The choice of the interpolation has some influence on the shape of criterion (1) over the transformation space, and hence on the behavior of the optimization algorithms used to match the images. Usually, the smoother the interpolation, the smoother the criterion and the fewer local minima [CVSM95, MCVM97]. However, this choice is independent of the matching algorithm itself, and so it will not be discussed here. In our implementations, we chose to use the trilinear interpolation.

Finally, images can be pre-processed before running a matching algorithm; for example, they can be smoothed by anisotropic diffusion or by a median filter. These pre-processings are usually intended to regularize the criterion to be minimized.

2.3 Notations for Iterative Non-Rigid Matching

Unless we constrain the transformation to belong to some parametric space, T is usually defined by its associated *displacement field* $\mathbf{U} = T - \mathrm{id}$ on the discrete image. While $T(\mathbf{p})$ represents the new coordinates of $\mathbf{p}$ after transformation, $\mathbf{U}(\mathbf{p}) = T(\mathbf{p}) - \mathbf{p}$ gives its displacement. In our implementation, vector fields are defined on every pixel of the image. If necessary, we use a trilinear interpolation to get a vector field defined everywhere in space.

Let us minimize iteratively the criterion (1) and denote $\bar{T}_n = \mathrm{id} + \bar{\mathbf{U}}_n$ the estimation of the solution at iteration n. To iterate from $\bar{\mathbf{U}}_n$ to $\bar{\mathbf{U}}_{n+1}$, we use a somewhat general 4-step algorithm:

1. We search a residual correction $\mathbf{u}_n$ of $\bar{\mathbf{U}}_n$.
2. We smooth $\mathbf{u}_n$ into $\bar{\mathbf{u}}_n$.
3. We add the smoothed correction to $\bar{\mathbf{U}}_n$, yielding $\mathbf{U}_{n+1} = \bar{\mathbf{U}}_n + \bar{\mathbf{u}}_n$.
4. We smooth this last displacement field into $\bar{\mathbf{U}}_{n+1}$, which is now the estimation of the solution at iteration $n+1$.

If we decide to smooth only total displacement fields $\mathbf{U}_n$, we have an elastic-like algorithm, since partial derivative equations of elastic models smooth the displacements of the elastic body. If we decide to smooth only correction fields $\mathbf{u}_n$, we have a viscous fluid-like approach, since partial derivative equations of viscous fluid models smooth velocities of the fluid body. Here, we assume that the $\mathbf{u}_n$ are small enough to be proportional to the velocities of the points of the image I evolving along iterations. Moreover, this terminology is correct only when filters are deduced from appropriate physical equations

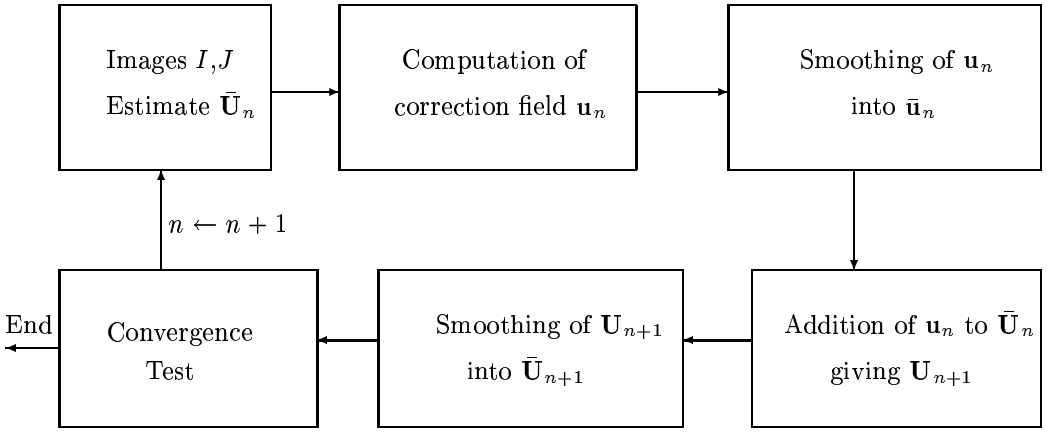


Figure 1: Overview of our 4-step iterative non-rigid matching algorithm

and are not generic low-pass filters like Gaussian filters. An example of such filters is established in [BNG96].

By default, the initial solution estimate $\bar{\mathbf{U}}_0$ is composed of zero vectors, so that $T_0 = Id$, but if we have an initial rigid transformation estimate we take its associated displacement field to initialize T_0.

The termination condition is performed on the correction field $\bar{\mathbf{u}}$: most of the time, such termination conditions can be written as $(N(\bar{\mathbf{u}}_n) < \epsilon)$, N being a norm on vector fields. So far, we have been unable to find a robust termination condition working on all types and sizes of images, and so the convergence decision is still manual.

In the following, we suppress the iteration index for easier reading in all previously introduced notation when it has no importance.

3 The "Demons"-Based Matching Algorithm

The "demons"-based method for gray-level images can be described by our general 4-step non-rigid matching method (fig. 1). It computes the correction field $\mathbf{u}$ with a method inspired by optical flow equations and does not smooth it before adding it to $\mathbf{U}$. $\mathbf{U}$ is then smoothed into $\bar{\mathbf{U}}$ by a Gaussian kernel, and the program is ready to iterate again.

3.1 Optical Flow

Suppose that we have a continuous set of images $I(t)$ of an objet $O(t)$ evolving over time, and that we want to retrieve the velocity of the physical points of $O(t)$ according to the intensity variation of the image pixels.

This is possible if we make the following fundamental assumption, which is a cornerstone of optical flow:

Intensity Conservation Assumption: We suppose that a physical point $\mathbf{p}(t)$ of $O(t)$ will always be observed with the same intensity in all images $I(t)$:

$$I(t)(\mathbf{p}(t)) = \text{constant} \tag{2}$$

If we differentiate equation (2) with respect to t, we get:

$$\nabla[I(t)](\mathbf{p})^T.\frac{d\mathbf{p}}{dt} + \frac{dI}{dt}(\mathbf{p}) = \nabla[I(t)](\mathbf{p})^T.\mathbf{v} + \frac{dI}{dt}(\mathbf{p}) = 0$$

where $\mathbf{v}$ is the velocity of the point $\mathbf{p}$ at time t. We thus get

$$\nabla[I(t)](\mathbf{p})^T.\mathbf{v} = -\frac{dI}{dt}(\mathbf{p})$$

This constraint on the velocity $\mathbf{v}$ is not enough to determine it completely. Only the projection of $\mathbf{v}$ on $\nabla[I(t)]$ is determined; this is the so-called aperture problem. Therefore, we also constrain $\mathbf{v}$ to be as small as possible. We then find

$$\mathbf{v} = -\frac{\frac{dI}{dt}(\mathbf{p})}{||\nabla[I(t)](\mathbf{p})||^2}\nabla[I(t)](\mathbf{p}) \tag{3}$$

3.2 From Optical Flow to "Demons"' Forces

Returning to our matching problem, we suppose that the object observed in I and J satisfies the intensity conservation assumption. This means that I and J should be of the same modality and should not be corrupted with an important bias or noise. Let us also make the following assumption:

Small Motion Assumption: We assume that the norm of displacement vectors between I and J are small[1].

In this case, we can approximate equation (3) at the first order. We can now write that the displacement of a point $\mathbf{p}$ between I and J is:

$$\mathbf{u}(\mathbf{p}) = \frac{I(\mathbf{p}) - J \circ \bar{T}(\mathbf{p})}{||\nabla I(\mathbf{p})||^2}\nabla I(\mathbf{p}) \tag{4}$$

This last equation can be seen as being the smallest displacement of $\mathbf{p}$ matching perfectly $J \circ \bar{T}(\mathbf{p})$ on a linear approximation of I at point $\mathbf{p}$ (see fig. 2).

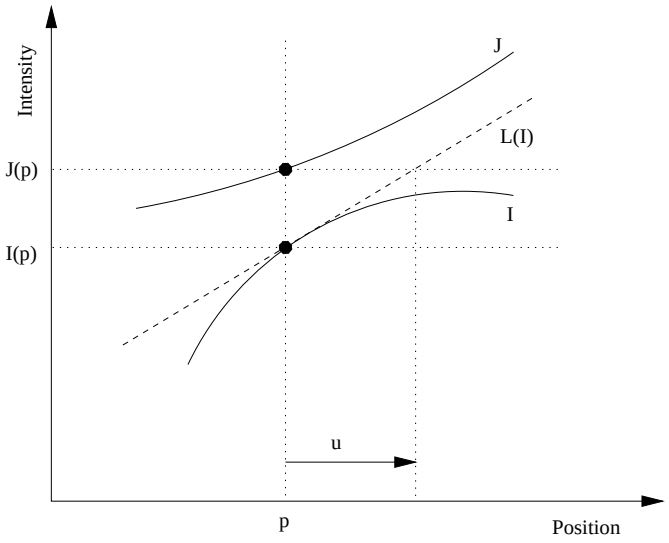


Figure 2: Point-to-point matching principle of formula (4) for a 1D image. $L(I)$ represents the linear approximation of I, and $\mathbf{u}$ the displacement found in formula (4).

[1] In practice, the notion of smallness here is relative to the size of the kernel used to compute derivatives on images, which is proportional to the distance between pixels.

However, formula (4) has the handicap of being unstable when the gradient of I become small. Thus, Thirion proposed to renormalize the equation into:

$$\mathbf{u}(\mathbf{p}) = \frac{I(\mathbf{p}) - J \circ \bar{T}(\mathbf{p})}{||\nabla I(\mathbf{p})||^2 + (I(\mathbf{p}) - J \circ \bar{T}(\mathbf{p}))^2} \nabla I(\mathbf{p}) \tag{5}$$

This last formula defines a motion correction field, which is called the "demons" forces in Thirion's terminology. Since this formula adds a variable homogeneous to an image intensity with a variable homogeneous to an image gradient, we believe we should instead consider the following equation:

$$\mathbf{u}(\mathbf{p}) = \frac{I(\mathbf{p}) - J \circ \bar{T}(\mathbf{p})}{||\nabla I(\mathbf{p})||^2 + \alpha^2 (I(\mathbf{p}) - J \circ \bar{T}(\mathbf{p}))^2} \nabla I(\mathbf{p}) \tag{6}$$

where α is a positive homogenization factor.

This renormalized displacement formula is bounded. Indeed,

$$||\mathbf{u}|| = \frac{|I - J \circ \bar{T}|.||\nabla I||}{||\nabla I||^2 + \alpha^2 (I - J \circ \bar{T})^2}$$

But $0 \leq (\alpha |I - J \circ \bar{T}| - ||\nabla I||)^2 = \alpha^2 (I - J \circ \bar{T})^2 + ||\nabla I||^2 - 2\alpha |I - J \circ \bar{T}|.||\nabla I||$, so displacements given by (6) are bounded[2] by $1/2\alpha$.

3.3 Transformation Smoothing

As discussed earlier, the transformation T should be kept smooth in order to regularize the minimization problem as well as to model, at least roughly, the physical constraints such as the continuity of the real transformation. However, nothing in formula (5) constrains T to be smooth, nor to preserve the topology of images. Equation (5) suggests the motion that each point is likely to have when considered separately; now, we would like to deduce the motion of the whole.

Usually, when the transformation is constrained to evolve in a low-dimensional parametric space, it is smooth *per se*. This is the case for rigid, affine or spline transformations. To retrieve the solution $\bar{T}$ from the displacement field $\mathbf{U}$, we only need to minimize a distance between them both. For example, if we are searching for the closest rigid transformation $(R^*, \mathbf{t}^*)$ in the sense of the SSD metric, we have to solve:

$$(R^*, \mathbf{t}^*) = \arg\min_{(R,\mathbf{t})} \sum_{\mathbf{p}} ||R(\mathbf{p}) + \mathbf{t} - (\mathbf{p} + \mathbf{U}(\mathbf{p}))||_2^2 \tag{7}$$

However, in the case of non-parametric, free-form transformations, a metric minimization alone is unconceivable because the solution will be very rough. Therefore, we have to add a smoothing term to the metric to be minimized. For example, we can try to solve the Tikhonov problem:

$$\arg\min_{\bar{\mathbf{U}}} \int (\bar{\mathbf{U}} - \mathbf{U})^2 + \int \sum_{\alpha \in \mathbf{N}^3} \frac{\sigma^n}{|\alpha|!} \left(\frac{\partial^{|\alpha|} \bar{\mathbf{U}}}{\partial^\alpha \mathbf{p}} \right)^2 \tag{8}$$

where $|\alpha| = \alpha_1 + \alpha_2 + \alpha_3$, for which the solution is known to be the smoothing of the displacement field $\mathbf{U}$ by a Gaussian kernel of standard deviation $4/\sigma^2$ [NFD94]. This last formulation draws our attention to the similarity between displacement fields used in registration and deformable models [Del94].

More generally, we can put the retrieval of $\bar{\mathbf{U}}$ in the framework of the smoothing and the warping of a vector field [Ara95], for which numerous techniques exist, ranging from simple linear smoothings

[2]The units in which are expressed the displacement vectors and consequently this limit are the same that of image gradient.

to complex non-linear physical models. For simplicity, Thirion chose to smooth $\mathbf{U}$ with a Gaussian, which happens to have a fast separable implementation for points lying on a grid:

$$\bar{\mathbf{U}}_n(\mathbf{p}_0) = \sum_{\mathbf{p}} \mathbf{U}_n(\mathbf{p}) e^{-\frac{||\mathbf{p}_0 - \mathbf{p}||^2}{2\sigma^2}} \tag{9}$$

4 Gradient Descent on a Matching Criterion

4.1 Sum of Square Differences Criterion Differentiation

Let us get back to the minimization problem (1) and take the SSD criterion as a metric between images:

$$d(I, J \circ \bar{T}) = C(\bar{T}) = \sum_{\mathbf{p}} (I(\mathbf{p}) - J \circ \bar{T}(\mathbf{p}))^2 \tag{10}$$

Now, let us move around $\bar{T}$ by adding a small correction field $\mathbf{u}$:

$$\begin{aligned} C(\bar{T} + \mathbf{u}) &= \sum_{\mathbf{p}} (I(\mathbf{p}) - J \circ (\bar{T} + \mathbf{u})(\mathbf{p}))^2 \\ &= \sum_{\mathbf{p}} \left[I(\mathbf{p}) - J \circ \bar{T}(\mathbf{p}) - \nabla J \circ \bar{T}^T.\mathbf{u}(\mathbf{p}) - \frac{1}{2}\mathbf{u}^T(\mathbf{p}).\mathcal{H}J \circ \bar{T}.\mathbf{u}(\mathbf{p}) \right]^2 \\ &= \sum_{\mathbf{p}} \left[I(\mathbf{p}) - J \circ \bar{T}(\mathbf{p}) \right]^2 + 2 \left[J \circ \bar{T}(\mathbf{p}) - I(\mathbf{p}) \right] \nabla J \circ \bar{T}^T.\mathbf{u}(\mathbf{p}) \\ &\quad + \left[\nabla J \circ \bar{T}^T.\mathbf{u}(\mathbf{p}) \right]^2 + \left[J \circ \bar{T}(\mathbf{p}) - I(\mathbf{p}) \right] \mathbf{u}^T(\mathbf{p}).\mathcal{H}J \circ \bar{T}.\mathbf{u}(\mathbf{p}) + o(||\mathbf{u}||_2^2) \end{aligned}$$

where $||\mathbf{u}||_2^2 = \sum_{\mathbf{p}} ||\mathbf{u}(\mathbf{p})||_2^2$, and $\mathcal{H}J$ is the Hessian of J. By identification, we obtain the gradient and Hessian of the criterion:

$$\nabla C(\bar{T}) = 2(J \circ \bar{T} - I)\nabla J \circ \bar{T} \tag{11}$$

$$\mathcal{H}C(\bar{T}) = 2(\nabla J \circ \bar{T}.\nabla J \circ \bar{T}^T + (J \circ \bar{T} - I)\mathcal{H}J \circ \bar{T}) \tag{12}$$

4.2 Second Order Gradient Descent

We now have the gradient and the Hessian of our SSD criterion in (12), which is all we need to minimize the criterion with a second order scheme.

For example, starting from $\bar{T}$, one minimization method consists in approximating the criterion by its tangential quadratic form at the current point $\bar{T}$:

$$C(\bar{T} + \mathbf{u}) \simeq C_{\text{quadra}}(\bar{T} + \mathbf{u}) = C(\bar{T}) + < \nabla C(\bar{T}) | \mathbf{u} > + \frac{1}{2}\mathbf{u}^T.\mathcal{H}C(\bar{T}).\mathbf{u}$$

If we suppose $\mathcal{H} > 0$, the minimum of this quadratic form is obtained where the gradient is the zero vector:

$$\nabla C_{\text{quadra}}(\bar{T} + \mathbf{u}) = \nabla C(\bar{T}) + \mathcal{H}C(\bar{T}).\mathbf{u} = 0 \qquad \Longleftrightarrow \qquad \mathbf{u} = -\mathcal{H}C(\bar{T})^{-1}.\nabla C(\bar{T})$$

If we replace in the previous formula the gradient and the Hessian of the criterion by their expressions (12), we obtain:

$$\mathbf{u} = -\left[\nabla J(\bar{T}).\nabla J(\bar{T})^T + (J(\bar{T}) - I)\mathcal{H}J(\bar{T}) \right]^{-1} (J(\bar{T}) - I)\nabla J(\bar{T}) \tag{13}$$

However, this last formula constrains the algorithm to invert a matrix for each pixel of the image. To accelerate the algorithm, we approximate the matrix to invert by its closest scalar matrix with respect to the norm:

$$||M||_2 = ||(m_{i,j})|| = \sqrt{\mathrm{tr}(M^T M)} = \sqrt{\sum_{i,j} m_{i,j}^2} \tag{14}$$

Now, for any matrix $M \in \mathcal{M}(n)$, we want to find the real number $\lambda(M)$ so that $\lambda(M) = \arg\min_{\lambda \in \mathbf{R}} ||M - \lambda.Id||_2$. Using the definition of the norm (14), we find that this boils down to minimizing $\sum_i (m_{i,i} - \lambda)^2$, and we thus have:

$$\lambda(M) = \frac{\mathrm{tr}(M)}{n}$$

So, finally, the scalar matrix closest to M with respect to the norm (14) is $(\mathrm{tr}(M)/n)I$.

Going back to our minimization problem and replacing the matrix to be inverted by its closest scalar matrix in expression (13), we obtain:

$$\mathbf{u} = -\frac{3(J(\bar{T}) - I)}{||\nabla J(\bar{T})||^2 + (J(\bar{T}) - I)\Delta J(\bar{T})} \nabla J(\bar{T}) \tag{15}$$

4.3 Explanation of "Demons" Forces

This last formula is strikingly similar to the formula (5) proposed by Thirion, for which the gradient of image I could be easily replaced by the gradient of image J if we inverse their roles in the argument we held in 3.2. However, since "demons" are motionless in the image I, it is more convenient to have a force proportional to ∇I, since this gradient won't have to be computed at each iteration. Also we better understand why we have an inversion in the role of I and J if we recall that a gradient descent on $d(I, J \circ T)$ matches points of I in J, while the "demons" algorithm matches points of J in I.

Let us determine the impact of this inversion on the gradient descent. If we invert the role of I and J, and minimize $d(I \circ T^{-1}, J)$, we find:

$$\mathbf{v} = -\frac{3(I(\bar{T}^{-1}) - J)}{||\nabla I(\bar{T}^{-1})||^2 + (I(\bar{T}^{-1}) - J)\Delta I(\bar{T}^{-1})} \nabla I(\bar{T}^{-1})$$

Please note that this time $\mathbf{v}$ is a correction field for $\bar{T}^{-1}$, so that our iterative formula is now $\bar{T}_{n+1}^{-1} = \bar{T}_n^{-1} + \mathbf{v}_n$. To get back to T, let us write

$$\mathbf{u}' = -\mathbf{v} \circ \bar{T} = \frac{3(I - J(\bar{T}))}{||\nabla I||^2 + (I - J(\bar{T}))\Delta I} \nabla I \tag{16}$$

We have $T_{n+1}^{-1} = \bar{T}_n^{-1} + \mathbf{v}_n = \bar{T}_n^{-1} - \mathbf{u}'_n \circ \bar{T}_n^{-1} = (\mathrm{id} - \mathbf{u}'_n) \circ \bar{T}_n^{-1}$, so

$$\begin{aligned} \bar{T}_{n+1} &= \bar{T}_n \circ (\mathrm{id} - \mathbf{u}'_n)^{-1} \\ &= \bar{T}_n \circ (\mathrm{id} + \mathbf{u}'_n + O(||\mathbf{u}'_n||_2^2)) \\ &= \bar{T}_n + (d\bar{T}_n)\mathbf{u}'_n + O(||\mathbf{u}'_n||_2^2)) \end{aligned}$$

So, when the differential $(d\bar{T}_n)$ of the transformation $\bar{T}_n$ is close enough to the identity and the corrections $\mathbf{u'_n}$ are small enough, inverting the role of I and J is then the same as inverting I and J in equation (15). Under these assumptions, we can consider that the "demons" algorithm for gray level images is an approximation of a second order gradient descent of an SSD criterion.

4.4 Differences between "Demons" and Gradient Descent

The "demons" formula (5) has a denominator which is the sum of two heterogeneous terms. Formulae (15) and (16) do not have this problem, since those two terms are homogenized by the introduction of a Laplacian. Moreover, unlike (5), these formulae commute with space similitudes, i.e. if we write $\bar{T}^{(I,J)}$ the transformation found to match I and J, then

$$S \circ \bar{T}^{(I \circ S, J \circ S)} = \bar{T}^{(I,J)} \circ S \tag{17}$$

where S is a similitude of factor s. Indeed, for the first iteration, $\bar{T}_0 = \text{id}$. In formula (15) we then have $\nabla(J \circ S) = S^T \nabla J \circ S$, so $||\nabla(J \circ S)||^2 = s^2 ||\nabla J \circ S||$, where $s = ||S||_2$. Similarly, $\mathcal{H}(J \circ S) = S^T(\mathcal{H}J \circ S)S$ so $\Delta(J \circ S) = s^2 \Delta J \circ S$. Finally, we have $s^2 \mathbf{u}^{(I \circ S, J \circ S)}(\mathbf{p}) = S^T \mathbf{u}^{(I,J)}(S\mathbf{p})$, but $S^T s^{-2} = S^{-1}$ so $S\mathbf{u}^{(I \circ S, J \circ S)}(\mathbf{p}) = \mathbf{u}^{(I,J)}(S\mathbf{p})$. If we assume that filters themselves commute with S, which is true for linear filters, formula (17) is true for $\bar{\mathbf{U}}_1$ and so it is also true for $\bar{T}_1$, and arguing the same way we prove by recurrence that it is true for $\bar{T}_n$ $\forall n$.

If we keep formula (13) and do not approximate its denominator, we obtain a matching algorithm that commutes with affine transformations:

$$A \circ \bar{T}^{(I \circ A, J \circ A)} = \bar{T}^{(I,J)} \circ A$$

Finally, we can note that all the formulae (5,13,15) are invariant over affine transformations on the intensities:

$$\bar{T}^{(A \circ I, A \circ J)} = \bar{T}^{(I,J)}$$

4.5 First Order Gradient Descent

Second order gradient descents have good convergence rates when they converge; however they may not converge if the Hessian of the function to be minimized behaves badly, like being near zero, negative, or badly conditioned. Moreover, medical images being noisy, particularly in the case of ultrasound images (see section 8.1), high order derivatives like the Hessian are more likely to be corrupt than lower order derivatives.

The computation of the function to be minimized is relatively fast, and the corrections $\mathbf{u}$ are small enough so that linear searches will not iterate very much anyway. Thus, we decided to use a first order gradient descent.

Some authors have already used a criterion gradient in their matching algorithms, but as far as we know, they all use a fixed fraction of this gradient and never make a gradient descent.

In this study, we have decided to begin with a very simple gradient descent, the linear search [PTVF92] along the steepest descent direction. In our case, this method gives good results.

In practice, we will not do a linear search along the whole gradient, because this would take too long. Instead, we minimize the criterion for each pixel independently, along the value of the gradient of the criterion at this pixel.

The algorithm is then:

1. Find the value of the descent direction $\nabla C(\bar{T})(\mathbf{p})$ of the criterion $C(\bar{T})$ at each point $\mathbf{p}$.
2. Do a linear search of the minimum for the criterion along the half-line starting from $\mathbf{p}$ and having direction $\nabla C(\bar{T})(\mathbf{p})$.
3. Correction field $\mathbf{u}$ is the result of this linear search.

The use of other gradient descent techniques is conceivable [Lem89, GMW81], especially the conjugate gradient method, which is a far better gradient descent method for complex, high-dimensional minimization problems. For a first approach, we believed this was not necessary since convergence is already fast with a steepest direction descent; however, this will be subject to further investigations.

5 Comparisons of Non-Rigid Matching Algorithms

In this section, we focus on a general method of comparison between non-rigid matching algorithms which will be used later in this report to evaluate our results.

Indeed, how to decide if the modifications made on a given non-rigid matching algorithm are an improvement or not? Reports on the performance of non-rigid matching algorithms are often bounded to eye comparisons or value of matching criterion. Although this may be sufficient for rigid matching, in non-rigid matching the smoothness of the displacement field is an equally important issue. Therefore, a comparison between two non-rigid matching algorithms cannot be based on the value of the minimized criterion only. Furthermore, if a non-rigid algorithm depends on a smoothing parameter, which is the case if T is not parametric, we should not compare algorithms with a single output but more *globally*.

Let us illustrate this with a trivial example. Let $A(\sigma)$ and $B(\sigma)$ be two non-rigid matching algorithms that depend on a smoothing parameter σ and that both minimize some common criterion, e.g., the SSD criterion. Suppose that after convergence, $B(\sigma)$ has always a better value for the criterion than $A(\sigma)$. Does it mean that B is superior to A? No. As an example, we could take $B(\sigma) = A(\sigma/2)$: obviously, B is not better than A. B will always find a better match, but its transformations will be always less smooth.

Because a non-rigid registration result has two parts, we propose to compare two algorithms in an image distance/transformation roughness space[3]. We take some distance d between images, which will quantify the matching quality of the algorithm. In the case where both algorithms minimize a criterion in which such a common distance appears, we should take this distance. We also take some measure of the roughness R on the displacement fields. Here again, if both algorithms smooth the displacement fields by minimizing some common roughness cost-function, this function should be taken. For an algorithm A that depends on a smoothing parameter σ, we measure the result of the registration for a particular σ using a two-dimensional vector $A(\sigma) = (d(A(\sigma)), R(A(\sigma)))$, composed of the distance between the matched images and the measure of roughness of the transformation.

In this distance-roughness space, we have a partial order relationship. We will say that $A(\sigma_1) > B(\sigma_2)$ if $d(A(\sigma_1)) < d(B(\sigma_2))$ **and** $R(A(\sigma_1)) < R(B(\sigma_2))$. Similarly, if $d(A(\sigma_1)) > d(B(\sigma_2))$ and $R(A(\sigma_1)) > R(B(\sigma_2))$ then $A(\sigma_1) < B(\sigma_2)$. In any other case, we *cannot* compare those results (see fig. 3).

As σ varies, the point $A(\sigma)$ moves and draws a curve in the distance-roughness space that we call the *solution curve* of the algorithm A (see fig. 4).

If the algorithm behaves properly, this solution curve should be decreasing, illustrating the trade-off between matching and smoothness of the displacement field, and bounded by its extremal points $A(0)$ and $A(\infty)$, respectively at the bottom right end and the top left end of the curve. In figure (4), the origin O represents the ideal rigid or affine[4] match. The point S, that should theoretically be one extremity of the solution curve of every non-rigid matching algorithm, represents the output of the best rigid or affine match, taken as an input of the non-rigid matching algorithm. The point L represents the limit of the algorithm, and usually does not lie on the y-axis either because the images cannot be perfectly matched, or because of the limited performance of the optimization algorithms.

Now, if the solution curve of B is above the solution curve of A, i.e. if for any solution $B(\sigma_2)$ there is a better solution given by $A(\sigma_1)$, then algorithm A is better than algorithm B.

In the following, we compare algorithms in an appropriate distance-roughness space. We have taken the SSD criterion as the distance between two images, since it is the criterion they minimize. As a roughness measure, we have taken:

[3] We use a distance-roughness space instead of a match-smoothness space, or any other combination, to stay coherent with our initial choice to see non-rigid registration and smoothing as a minimization, not as a maximization.

[4] depending on the roughness measure

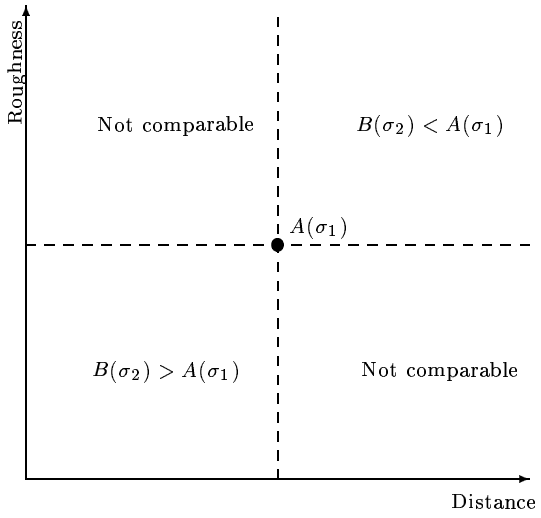


Figure 3: Order relationship in a distance-roughness space. If $B(\sigma_2)$ lies in the bottom left quarter plane, it is superior to $A(\sigma_1)$. If it lies in the up right quarter plane, it is inferior to $A(\sigma_1)$. If it lies elsewhere, both points are *not comparable*.

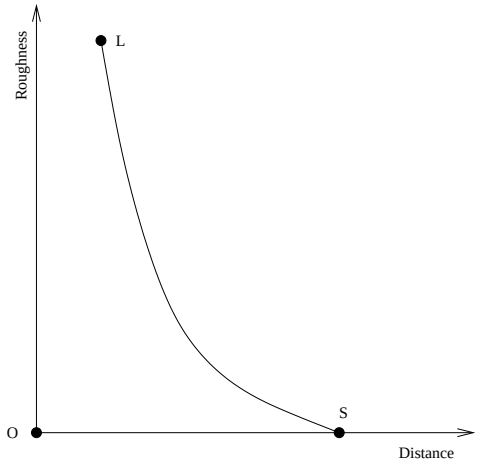


Figure 4: Solution curve of a non-rigid matching algorithm in a distance-roughness space. O represents the (impossible) ideal rigid match. S represents the best rigid match possible, which is one extremity of the solution curve. L is the other extremity and represents its limited capacity to minimize the distance between images.

$$\int_{\mathbf{R}^3} ||d\bar{\mathbf{U}}||_2^2 = \int_{\mathbf{R}^3} ||\partial_1 \bar{\mathbf{U}}||_2^2 + ||\partial_2 \bar{\mathbf{U}}||_2^2 + ||\partial_3 \bar{\mathbf{U}}||_2^2 \tag{18}$$

which is the first order approximate of the smoothing term of the Tikhonov problem (8), of which Gaussian smoothing is the closed-form solution.

6 Introducing Smoothing Weights

6.1 *A priori* Weights

As the image gradient tends toward 0, its direction becomes less and less robust and consequently its informative value becomes lower and lower, and when the gradient is actually null, it is useless for matching. In the extreme case of optical flow, where we had a blind confidence in the linear approximation of I given by the gradient, this results in an unstable formula for points with low gradient norms, which we therefore had to renormalize (turning (4) into (5)). Even in the case of a linear search, which is free from numerical explosion no matter how small the gradient is, taking into account the relative informative values of the direction of search seems wise, and we should especially avoid taking into consideration a match found on points where the gradient is so small that the direction of the search is random.

Previously, for the "demons" algorithm, an attempt to do so has been done by introducing a multiplicative factor on vectors of $\mathbf{u}$ which tends to 0 along with the norm of the gradient. This means that on those points where the gradient is almost null, the displacement will be null. Now, we are not concerned if those points do not move as long as they have *no impact* during the computation of $\bar{\mathbf{u}}$ from $\mathbf{u}$.

Therefore, we propose to drop multiplicative factors, and to replace them by *weights*. We will link each pixel $\mathbf{p}$ of the image to a weight $\omega(\mathbf{p})$ which will be increasing along with the norm of the gradient at this point. Among other choices, we could conceive to use:

1. $\omega(\mathbf{p}) = \mathbf{1}_{||\nabla J \circ T(\mathbf{p})|| > \alpha}$ (threshold)
2. $\omega(\mathbf{p}) = 1 - e^{-\alpha||\nabla J \circ T(\mathbf{p})||}$ (bounded weights)
3. $\omega(\mathbf{p}) = ||\nabla J \circ T(\mathbf{p})||$ (unbounded weights)

The coefficient α of weights (1) and (2) should be chosen according to image the I, e.g., by taking as α some quantile of the distribution of $||\nabla J \circ T||$ on the image. Here, however, we will present results with weight (3), since it seems to give better results on the matching situations we tested, while has the advantage of being parameterless. Furthermore, we introduced an *a priori* weight in the "fluid" filtering step, i.e. when we retrieve $\bar{\mathbf{u}}$ from $\mathbf{u}$, and let the "elastic" filtering be non-weighted Gaussian, since *a priori* weights were conceived to discriminate insignificant results of the linear search.

These weights are introduced very naturally into the criteria used to compute $\bar{\mathbf{u}}$ from $\mathbf{u}$ or $\bar{\mathbf{U}}$ from $\mathbf{U}$ (section 3.3): they are simply multiplicative factors of the matching part of the criteria. For example, if we are searching for a rigid transformation from the displacement field $\mathbf{U}$ with a weighted SSD criterion, equation (7) is replaced by:

$$(R^*, \mathbf{t}^*) = \arg\min_{(R,\mathbf{t})} \sum_{\mathbf{p}} \omega(\mathbf{p})^2 ||R(\mathbf{p}) + \mathbf{t} - T(\mathbf{p})||_2^2 \tag{19}$$

In the same way, for a non-rigid matching, the equation (8) is replaced by:

$$\arg\min_{\bar{\mathbf{U}}} \int \omega(\mathbf{p})^2 (\bar{\mathbf{U}} - \mathbf{U})^2 + \int \sum_{\alpha \in \mathbf{N}^3} \frac{\sigma^n}{|\alpha|!} \left(\frac{\partial^{|\alpha|} \bar{\mathbf{U}}}{\partial^\alpha \mathbf{p}} \right)^2 \tag{20}$$

General directions to introduce weights in other non-variational warping methods used in non-rigid matching include the fact that they must not be simple multiplicative factor of the filter and in particular that a renormalization must occur; for example, the Gaussian smoothing (9) is replaced by:

$$\bar{\mathbf{u}}(\mathbf{p}_0) = \frac{\sum_{\mathbf{p}} \omega(\mathbf{p})\mathbf{u}(\mathbf{p}) e^{-\frac{||\mathbf{p}_0-\mathbf{p}||^2}{2\sigma^2}}}{\sum_{\mathbf{p}} \omega(\mathbf{p}) e^{-\frac{||\mathbf{p}_0-\mathbf{p}||^2}{2\sigma^2}}} \tag{21}$$

Any other linear filter can be weighted the same way.

The implementation of a weighted linear filter is easy, and in the case of the weighted Gaussian filter, we keep the separability property, which means low computation times. To filter some vector field $\mathbf{u}$ with associated weights $\omega(\mathbf{u})$, we just have to:

1. Create the field $\omega.\mathbf{u}$, where all vectors have been multiplied by their weight.
2. Smooth ω and $\omega.\mathbf{u}$ with the filter you want to weight[5].
3. Divide the second by the first.

To illustrate the interest of this method, we first show an exemple of registration between two circles (fig. 5). These images are composed mainly of pixels with null derivatives, lots of them being already in correspondance with a pixel of its same color (namely, black background pixels and the center of the circle). Only a small number of pixels want to move. Without *a priori* filtering, those pixels will have difficulty to move alone the whole image, and so convergence is quite slow. That's why, at iteration 20, the solution curve of the algorithm using *a priori* filtering is clearly superior to the other (see fig. 5(c)). When convergence has occurred (iteration 100), using weighted smoothing is still superior to regular smoothing for half of the curve. However, *a priori* smoothing, because of its renormalization, is unable to find displacement fields as smooth as Gaussian smoothing, and so Gaussian smoothing is superior for high values of the smoothing parameter (where roughness measure is low).

We also present results on a 2D IRM, which has been synthetically deformed (fig. 6). On this more realistic example, where areas with no informative points are smaller and therefore easier to move, the improvement in speed is less obvious. In the other hand, because informative points are now dense in the image, transformations obtained with *a priori* filtering are now able to be as smooth and actually smoothier than Gaussian smoothing, yielding a solution curve at convergence (iteration 20) that is better everywhere (see fig. 6(c) and 6(d)).

6.2 *A posteriori* Weights

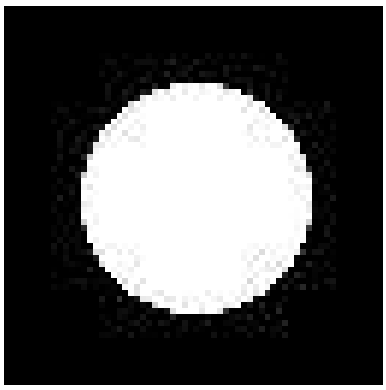
The previous weighted filter considers that results on points where the gradient norm is low are insignificant. Indeed, we know that on these points, the search direction is random.

On the other hand, once the linear search is done and $\mathbf{U}$ is found, we have an estimate of the solution. Now, we can judge the results on this match and thus give importance to the points where the results are good. The *a posteriori* weights for getting $\bar{\mathbf{U}}$ from $\mathbf{U}$ implement this idea.

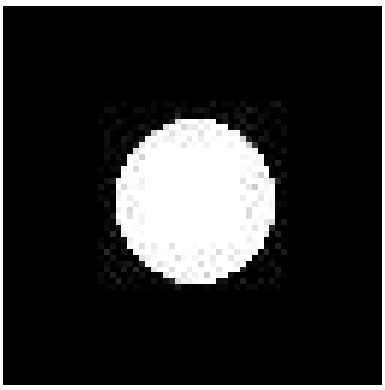
Please note that *a posteriori* weighting makes sense because we are doing a linear search, and are sure to be at a local minima. This weighting does not work with "demons" forces, or even worse, with a fixed fraction of the gradient, because during the first iteration, results are bad almost everywhere, and the discrimination done by the filter is meaningless.

The theoretical difficulty with *a posteriori* weights is that it should be a measure of the fit of a couple of points, while the metric d we choose is defined on the whole image. For some metrics, like SSD, correlation coefficient, or correlation ratio, the deduction of a metric on the points from the

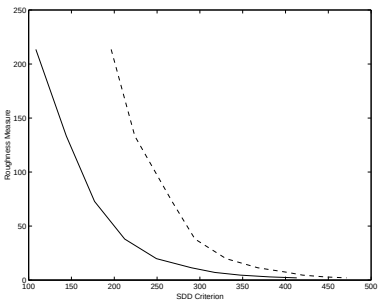
[5] Here, we can easily deduce a filter applying to a scalar field (ω) from a filter applying to a vector field ($\omega.\mathbf{u}$) because it is linear.



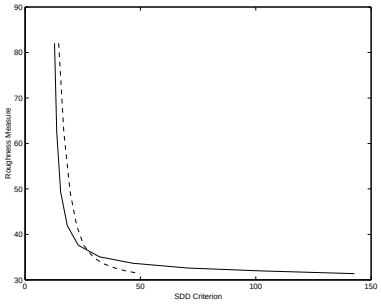
(a) Reference image



(b) Study image



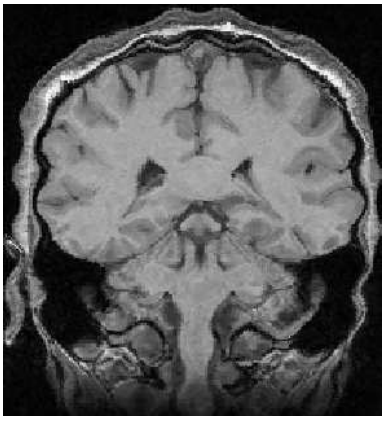
(c) Solution curve at iteration 20



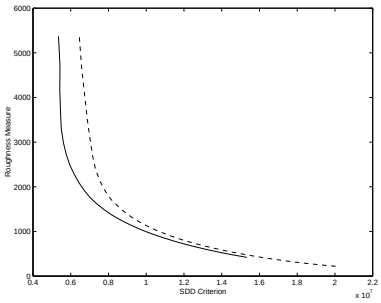
(d) Solution curve at convergence

Figure 5: Two circles to match and the solution curves for both versions of the algorithm. Plain curve: weighted Gaussian filtering. Dashed curve: standard Gaussian filtering.

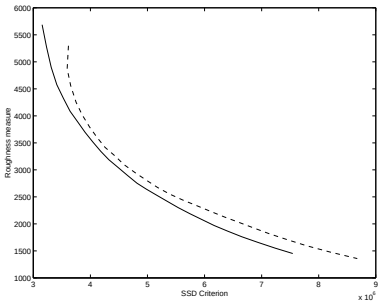
(a) Reference image



(b) Study image



(c) Solution curve at iteration 5



(d) Solution curve at convergence

Figure 6: Two 2D MRI to match and the solution curves for both version of the algorithm. Plain curve: weighted Gaussian filtering. Dashed curve: standard Gaussian filtering.

metric on the image is natural because there is an underlying statistical model, for which the white noise assumption model has been made, which means that these metrics are in fact sums of a metric on all couple of points.

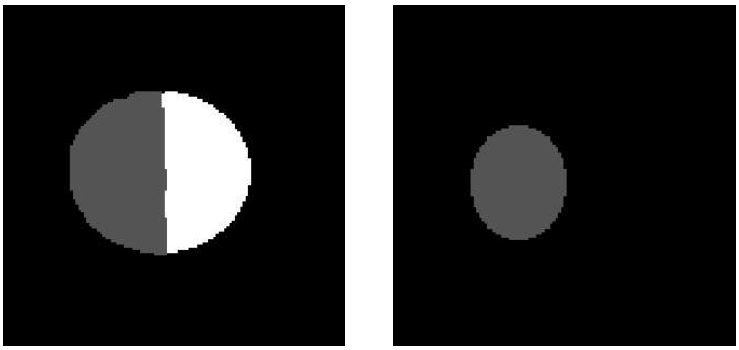
For our SSD criterion, we can choose for example:

1. $\omega(\mathbf{p}) = \mathbf{1}_{(I(\mathbf{p})-J\circ T(\mathbf{p}))^2<\beta}$ (threshold)
2. $\omega(\mathbf{p}) = 1/(1+\beta(I(\mathbf{p})-J\circ T(\mathbf{p}))^2)$ (low decrease)
3. $\omega(\mathbf{p}) = e^{-\beta(I(\mathbf{p})-J\circ T(\mathbf{p}))^2}$ (fast decrease)

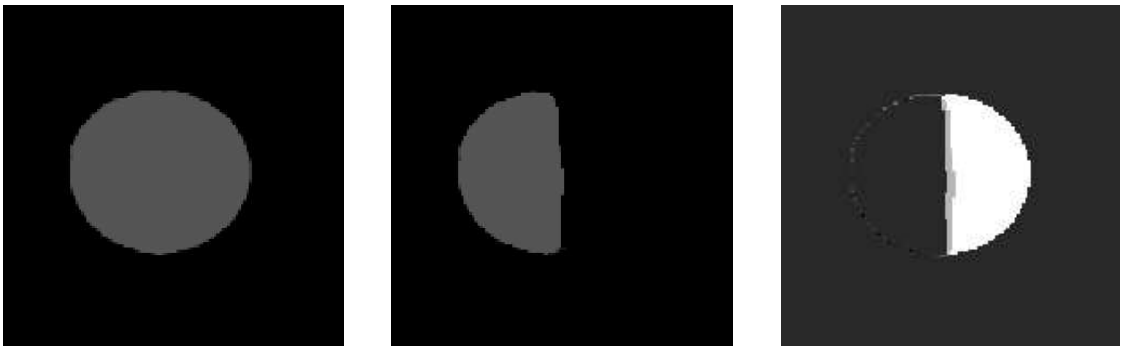
Here, we present results using the third kind of weight.

This kind of weighting enables us to tackle partially body appearance or disappearance. The distinction between the intensity range of two bodies is done through β, which we can conceive for further developments as being a function of intensity.

To show its properties, we tested it on an artificial example. Figure 7 presents two images to match. The reference image, on the left, contains two connected areas, one gray and one white. The study image, on the right, only has a single gray area.



(a) Reference image (b) Study image



(c) Registration without *a posteriori* smoothing (d) Registration with *a posteriori* smoothing (e) Subtraction of 7(b) to 7(a)

Figure 7: First row, two images to register, with white intensity matter apparition. Second row, registration results.

When the images 7(a) and 7(b) are matched without *a posteriori*-weighted smoothing, the gray ellipse of image 7(b) is stretched in order to fully recover the bicolored ellipse of 7(a). Indeed, it is

really this position that minimizes the mean square difference between both images. If we introduce *a posteriori* weights, the gray ellipse stretches only in the gray area of 7(a). However, the gray ellipse still overlaps a small border of the white area, because an *a posteriori* smoothing cannot totally undo the tendency of recovering the whole bicolored ellipse of the correction field, which is still a gradient descent of the SSD criterion.

This explains why we cannot compare *a posteriori* filters with basic gaussian filtering in some distance-roughness, because we do not have a common distance minimized by both algorithms. They are not doing the same job, and therefore, the choice of using *a posteriori* weights or not depends on the problem, like the presence of occlusion or matter (dis)appearance.

7 Multiresolution Approach

The descent gradient approach assumes that the displacements that has occurred between the acquisitions of the two images to be matched are small. Indeed, it is well known that gradient descents are very easily stuck in local minima. Even if smoothing the displacement fields regularizes the problem, this is still not sufficient.

In order to avoid local minima, we could use another minimization algorithm that would be less sensitive to this issue, e.g. simulated annealing or a genetic algorithm [LLVCG96]. However, computation time seriously increases with these techniques. In imagery, a classical technique improves the size of the convergence area without changing our minimization tool: multiresolution.

Given an image I, multiresolution consists in constructing a set of images $I^{(k)}$, which size decreases as k increases. When two images I and J have been reduced, the regions are all close to each other, and the assumption of small displacement reasonably holds.

However, the reduced images $I^{(k)}$ and $J^{(k)}$ are much less detailed than original images I and J (which are incidentally $I^{(0)}$ and $J^{(0)}$), so the solution $\bar{T}^{(I^{(k)},J^{(k)})}$ found when matching this pair of images will not be very precise. Thus, once $I^{(k)}$ and $J^{(k)}$ registered, we want to register $I^{(k-1)}$ and $J^{(k-1)}$, with $\bar{T}^{(I^{(k)},J^{(k)})}$ as an initial estimate. Here, the assumption of small displacement also reasonably holds, because the corrections needed by $\bar{T}^{(I^{(k)},J^{(k)})}$ are due to the difference in the detail level between $(I^{(k)}, J^{(k)})$ and $(I^{(k-1)}, J^{(k-1)})$, which is not important. Also, because images $(I^{(k-1)}, J^{(k-1)})$ are bigger, $\bar{T}^{(I^{(k)},J^{(k)})}$ has to be oversampled to fit them. We used a trilinear interpolation for this.

To create a multiresolution pyramid, we have three steps:

1. The first reduced image is obtained by subsampling the original image. All dimensions of this subsampled image are a power of two, which are the largest strictly inferior to the size of the image. Trilinear interpolation is used for subsampling.

2. Further reductions are obtained by recurrence. From the previous image $I^{(k)}$, we construct $I^{(k+1)}$ so that the dimensions of $I^{(k+1)}$ are half those of $I^{(k)}$, and each pixel of $I^{(k+1)}$ is the mean of the 8 pixels it replaces.

3. We stop iterating when the dimensions of the images are below some value.

8 Results

8.1 Registration of 3D Ultrasonic Images

In this experience, we have register two 3D ultrasound images provided by the European project ROBOSCOPE. A catheter has been introduced into the brain of a dead pig, and a balloon located at its end has been inflated to various volumes, in order to simulate matter displacement that occurs around a tumor during excision. A 3D ultrasound image is taken at various volumes.

Images presented in figure 8 are isotropic sub-images of size $75 \times 62 \times 73$ pixels extracted from the original images, and centered on the balloon. Here, the difficulty for matching comes from the strong speckle noise that corrupts ultrasound images. This noise perturb the original contours (i.e., gradient of the image). This generally hinders the good behavior of iconic methods, and in any case makes it impossible to retrieve a transformation T with a great accuracy. T should therefore be smoothed more than usual, not only to average then displacement where noise is important, but also to propagate the displacement where no information is to be found, as in this case where a whole section of the balloon (lower left part) is not visible.

Here, the matching has been done on low image resolution. This constrains the transformation to be very smooth, which is advisable for ultrasound images. Registration is therefore fast (40.3 sec. on an Alpha PWS 500MHz workstation).

The registration result is visually correct. However, to make sure that the displacement field depicted in figure 8(c) is also correct, we have computed its Jacobian ($= |d\bar{\mathbf{U}}|$) with David Rey's tools, initially aimed to automate the detection and the segmentation of evolving brain lesions [RSDA98]. An important result is the smoothness of the Jacobian isovalues. In particular, the fact that one Jacobian isovalue follows the outline of the balloon and almost segments it tends to prove the displacement field we found has a physical meaning. It has some problems following the bottom left part of the outline because that part of the balloon is missing in those images, and because the left part of the balloon does not move since the catheter was fixed.

8.2 Registration of Lung 3D Scanner Images

In this experience, we had to match two scanner images of the same patient, who suffers from a melanoma that has evolved over time, provided by Pr. Grenier of the Pitié-Salpétrière hospital in Paris. Here, the difficulty arises from the occlusion that hides part of the melanoma in both images. Of course this occlusion is different between both images (see images 9(a) and 9(b)).

An initial rigid matching has been done thanks to Sébastien Ourselin's tools [ORSP98]. Then, we have registered those images with and without *a posteriori* smoothing. Without *a posteriori* smoothing, the left side of the melanoma of image 9(a) is compressed in order to match the shape of the melanoma in image 9(b), as shown in image 9(c). These unrealistic displacements (image 9(e)) can be prejudicial for displacement field analysis based algorithms, e.g. for the evolution estimation. With *a posteriori* smoothing, the left side of the melanoma has moved less (image 9(d)), and the occlusions have less impact on displacement field (image 9(f)).

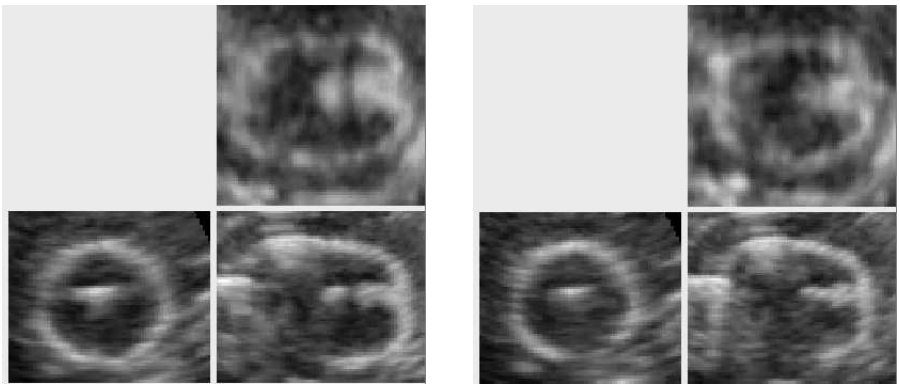
9 Conclusion and Future Work

The diffusion, or "demons"-based, approach for non rigid matching consisted in computing a displacement field by smoothing a "force" field defined in an *ad hoc* way. Here, we have preferred to place non rigid matching in a criterion minimization framework. We have chosen a simple criterion, the SSD criterion, corresponding to a simple assumption, image intensity conservation.

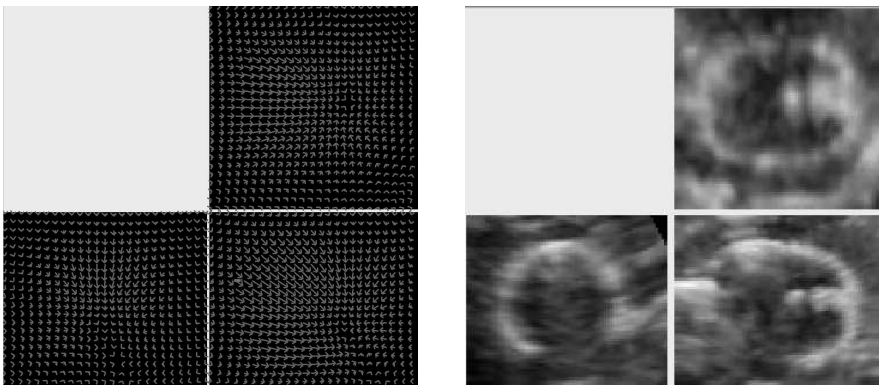
Then, we used a gradient descent, coupled with a multiresolution approach, to minimize this criteria. We proved that the "forces" used by Thirion for his "demons" are an approximation of a second order gradient descent scheme. This theoretically explains and justifies this until now heuristic algorithm. In our implementation, we preferred to use a first order gradient descent, arguably better for registration purposes.

During the criterion minimization step, we have introduced two novel weighted smoothings:

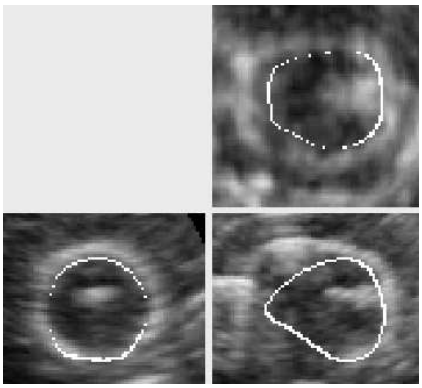
1. the *a priori* smoothing is applied to correction fields. It takes into account the confidence we can have on results obtained at each pixel, giving more weight to pixels with a higher gradient norm. The *a priori* smoothing thus helps the gradient descent to find a better solution.



(a) Study image (b) Reference image

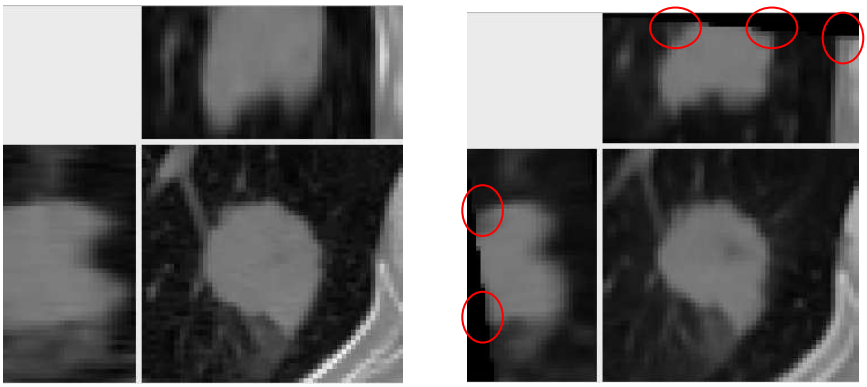


(c) Displacement field (d) Registered image



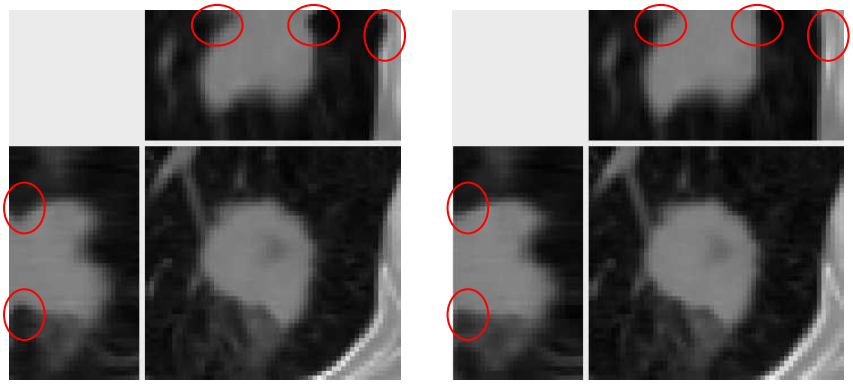
(e) A Jacobian isovalue

Figure 8: Registration results on ultrasound images. Images were provided by Paul Volker, Fraunhofer (Institute) Society, Munich (Germany) as part of the EC-funded ROBOSCOPE project.



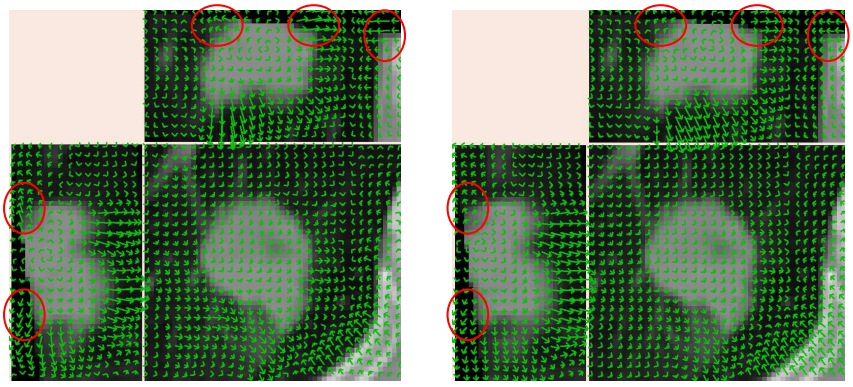
(a) Image to match

(b) Reference image



(c) Registration result without *a posteriori* smoothing

(d) Registration result with *a posteriori* smoothing



(e) Displacement field found without *a posteriori* smoothing

(f) Displacement field found with *a posteriori* smoothing

Figure 9: First row: Two scanner images of a melanoma with occlusion. Middle row: Registration results with and without *a posteriori* smoothing. Last row: Associated displacement fields. Area improved by *a posteriori* smoothing are stressed in circles.

2. the *a posteriori* smoothing is applied to displacement fields. It takes into account the accuracy of the matching at each point, and corrects bad matches with displacements obtained on correctly matched neighboring points. The *a posteriori* is thus able to tackle matter (dis)appearance and occlusion problems, which are particularly sensitive issues for iconic non-rigid methods.

Finally, we proposed a rigorous framework to compare some non-rigid matching algorithm, distance-roughness spaces, which take into account both the matching quality and the smoothness of results.

Following this work, we see four essential research directions:

1. Minimization criterion, as well as *a priori* weights, could be deduced from statistical assumptions;
2. The registration and smoothing steps should be melted into a single registration-and-smoothing step, by means of minimizing a single criterion which explicitly takes into account matching and smoothing, as commonly done for active contours;
3. The multiresolution step should also be explicitly integrated into the minimization framework;
4. A truly automatic termination condition, working no matter the nature and size of the images to match, still has to be found.

From an algorithmic point of view, this work has numerous potential extensions. First, as *a priori* and *a posteriori* are independent of the filter we apply to vector fields, it is possible to replace Gaussian smoothing by other filters, e.g. linear elastic filters. Then, developments done here for the SSD criterion can be done for other criteria, and especially for multimodal criteria. However we should expect greater complexity and longer computation time for non-local criteria. Finally, using more complex gradient descents and comparing them should enable us to make the best choice possible between complexity and computation time.

Acknowledgments

We would like to thank Pr. Grenier from the Pitié-Salpétrière hospital in Paris for providing the lung scanner images. We would like to thank Paul Volker, Fraunhofer (Institute) Society, Munich (Germany) who provided us with the US images as part of the EC-funded ROBOSCOPE project HC 4018, a collaboration between The Fraunhofer (Institute) Society (Germany), Fokker Control System (Netherlands), Imperial College (U.K.), INRIA Sophia Antipolis (France), ISM Salzburg, and Kretz Technik (Austria). Thanks also to Kretz and ISM for providing a high-end imaging system for these examinations.

Thanks very much to Janet Bertot for the proofreading of this article.

References

[Ara95] N. Arad. *Image Warp Design based on Variational Principles.* Ph.D. dissertation, Tel-Aviv University, January 1995.

[BK89] R. Bajcsy and S. Kovačič. Multiresolution elastic matching. *Computer Vision, Graphics and Image Processing*, 46:1–21, 1989.

[BNG96] M. Bro-Nielsen and C. Gramkow. Fast fluid registration of medical images. In *Proceedings VBC'96*, pages 267–276, 1996.

[Bro81] C. Broit. Optimal registration of deformed images. August 1981. Doctoral Dissertation, University of Pennsylvania.

[BT88] J. C. Bamber and M. Tristam. *The physics of Medical Imaging*, chapter 7, pages 319–388. *Medical Science Series*, Adam Hilger, 1988.

[CJM97] G. E. Christensen, S. C. Joshi, and M. I. Miller. Volumetric transformation of brain anatomy. *IEEE Transactions on Medical Imaging*, 16(6):864–877, December 1997.

[CVSM95] A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal. 3d multi-modality medical image registration using feature space clustering. In N. Ayache, editor, *Proc. of the First Int. Conf. on Comp. Vision, Virtual Reality and Robotics in Medicine (CVRMed'95)*, pages 195–204, Springer Verlag, 1995.

[Del94] H. Delingette. Simplex meshes: a general representation for 3d shape reconstruction. In *Proc. of Int. Conf. on Computer Vision and Pattern Recognition (CVPR'94)*, Seattle, USA, June 1994.

[GMW81] P. E. Gill, W. Murray, and M. H. Wright. *Practical Optimization.* Academic Press, 1981.

[Lem89] C. Lemarechal. *Méthodes Numériques d'Optimisation. Collection Didactique*, INRIA, 1989.

[LJ94] L. Lemieux and R. Jagoe. Effect of fiducial marker localization on stereotactic target coordinate calculation in ct slices and radiographs. *Phys. Med. Biol.*, 39:1915 – 1928, 1994.

[LLVCG96] E. Lutton, J. Lévy Véhel, G. Cretin, P. Glevarec, and C. Roll. *Mixed IFS: Resolution of the Inverse Problem Using Genetic Programming.* Technical Report RR-2631, INRIA, 1996.

[MCVM97] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. In *IEEE Transactions on Medical Imaging*, pages 187–198, April 1997.

[MMV98] J. B. A. Maintz, E. H. W. Meijering, and M. A. Viergever. General multimodal elastic registration based on mutual information. *Image Processing*, 1998.

[MV98] J. B. A. Maintz and M. A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.

[NFD94] M. Nielsen, L. Florack, and R. Deriche. *Regularization and Scale Space.* Technical Report RR-2352, INRIA, September 1994.

[ORSP98] S. Ourselin, A. Roche, G. Subsol, X. Pennec, and C. Sattonnet. *Automatic Alignment of Histological Sections for 3D Reconstruction and Analysis.* Technical Report RR-3595, INRIA, December 1998.

[PTVF92] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipies in C.* Cambridge University Press, 1992.

[RMPA98] A. Roche, G. Malandain, X. Pennec, and N. Ayache. *Multimodal Image Registration by Maximization of the Correlation Ratio.* Technical Report RR-3378, INRIA, August 1998.

[RSDA98] D. Rey, G. Subsol, H. Delingette, and N. Ayache. *Automatic Detection and Segmentation of Evolving Processes in 3D Medical Images: Application to Multiple Sclerosis.* Technical Report RR-3559, INRIA, November 1998. Available at http://www.inria.fr/RRRT/RR-3559.html.

[SHH95] C. Studholme, D. L. G. Hill, and D. J. Hawkes. *Information Processing in Medical Imaging*, chapter Multiresolution Voxel Similarity Measures for MR-PET Registration, pages 287–298. Kluwer Academic Publishers, 1995.

[Sub98] G. Subsol. Crest lines for curve based warping. In A. W. Toga, editor, *Brain Warping*, chapter 14, pages 241 – 262, Academic Press, October 1998.

[Thi98] J.-P. Thirion. Image matching as a diffusion process: an analogy with maxwell's demons. *Medical Image Analysis*, 2(3), 1998.

[VHSL98] B. C. Vemuri, S. Huang, S. Sahni, C. M. Leonard, C. Mohr, R. Gilmore, and J. Fitzsimmons. An efficient motion estimator with application to medical image registration. *Medical Image Analysis*, 2(1):79–98, 1998.

[Vio95] P. A. Viola. *Alignment by Maximization of Mutual Information.* Ph.D. dissertation, M.I.T. Artificial Intelligence Laboratory, ftp://publications.ai.mit.edu, 1995.

[WKJK98] S. K. Warfield, M. Kaus, F. A. Jolesz, and R. Kikinis. Adaptive template moderated spatially varying statistical clas sification. In *Proc. of MICCAI'98*, pages 431–438, Cambridge, USA, 1998.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - B.P. 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 Villers lès Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot St Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, B.P. 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399